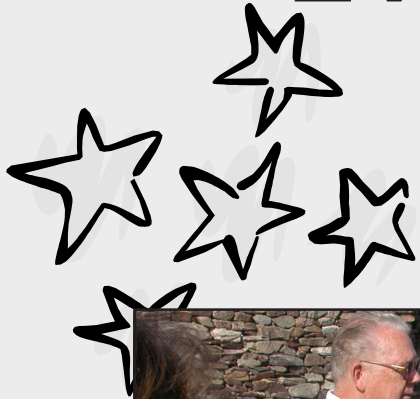# Evaluation

Having developed the art of interpretation to a high degree, (we are) increasingly stepping back and looking at the science of evaluating interpretation.

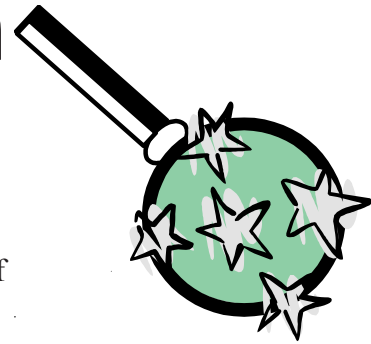William Penn Mott, Jr.

© 2003 California State Parks

Evaluation

What is it? planned assessment

Why do we do it? to determine effectiveness of our interpretive efforts

How do we do it? with systematic examinations of interpretive programs

> **There is much to be learned from evaluating interpretation, if we only ask the right questions.**

Introduction

Evaluation (assessment) is a critical component of any interpretation. It links many of the pieces that we have discussed in other modules. Evaluation provides feedback in the communication process (Module 3-Communication), continues the planning cycle (Module 4-Planning), measures the benefits and values of programs (Module 2-Purpose and Values), and helps set the direction for the future (Module 13-Professionalism). This module will review the basic evaluation process, the purpose of evaluation, and the targets and sources of evaluation. Finally, it will examine some practical methods of conducting evaluation. It should be noted that while this module will focus on personal interpretation, nonpersonal interpretation (exhibits, displays, brochures, etc.) should also be regularly evaluated for effectiveness.

This module is intended to be used in conjunction with the *Aiming for Excellence* evaluation handbook (McDonald, 2000), which is a resource providing guidelines and methods (complete with forms) of evaluation in California State Parks. We will focus on evaluation conducted by the interpreter with little to no experience in it. It is meant to function as an overview and provide useful, practical supplements to other evaluation materials provided by the park and the Department. The main thing to remember is that conducting meaningful evaluation is generally up to you, the interpreter. It is your responsibility to seek out assessment of your own programs. The value of your work will be directly affected by the quality of your evaluation efforts.

Before addressing practical methods of evaluation, let us begin by looking at the basic concepts and theories of evaluating interpretation. The following are the guiding questions that should be used to design these procedures: *Why conduct evaluation? What should be evaluated? When should evaluation be conducted? Who is conducting the evaluation?* and finally*, How should evaluation be conducted?* The answers to these questions directly impact which methods and procedures of evaluation are appropriate. Let us review the potential answers to each of these questions and discuss the resulting effects.

> *All interpretive staff who present interpretive programs will have their program(s) evaluated a minimum of twice per year, including a minimum of one evaluation by an interpretive coordinator, lead person or supervisor using the Standard RAPPORT Evaluation DPR 461 form. Additional evaluations may be conducted using a variety of appropriate techniques.*
>
> Department of Parks and Recreation Policies, Aiming for Excellence

Why

There are many purposes and benefits of conducting assessment. Each is important and provides insight into the overall success of interpretation. One of the main purposes is to increase program effectiveness. Reflecting on Module 3-Communications there are many elements that determine and influence overall program effectiveness. As a result, effectiveness should be assessed using several different indicators.

The first question of evaluation is to determine the purpose of the assessment. You must know what the goals and objectives of the evaluation are *before* you can design the appropriate strategy. For example, if you are trying to assess the cost-effectiveness of programs, asking visitors how satisfied they are with programs will not yield appropriate or useful data. Meeting visitors' needs, addressing management goals and objectives, and communicating effectively are all individual elements that should be used to indicate program success. Evaluation measures, properly implemented, can indicate which specific areas need to be improved to increase overall program effectiveness. In Module 2-Purpose and Values we learned that there are numerous reasons for interpretive services. Evaluation will assist in highlighting successes and promoting improvements. Without effective evaluation it would be difficult to know what is working, for whom, and why.

Another key purpose of conducting evaluation is to improve your personal skills as an interpreter. Evaluation can help you *objectively* assess your verbal and nonverbal communication skills and your ability to research, organize, and present a program.

Evaluation should also be conducted to assess whether the goals and objectives of the program have been met. Programs should have been carefully created to address specific goals and objectives (see Module 4-Planning). Assessment allows for a critical analysis of whether you have met those targeted objectives–you want to assess your program's ability to do what it set out to do.

Being able to justify budgets, measure cost-effectiveness, and demonstrate the financial benefits of interpretation is another positive aspect of evaluation, especially for administrators. In tight budgetary times, the ability to provide cost-benefit analysis of interpretive services is highly productive.

It should be clear from the other modules in this handbook that overall effectiveness of interpretive programs is not a reflection of any one specific element–effectiveness is determined by a combination of characteristics.

Why conduct evaluation?

- **Increase program effectiveness.**
- **Determine if visitors' needs are met.**
- **Assess if program goals and objectives are met.**
- **Assess interpreter's effectiveness.**
- **Establish professional legitimacy.**
- **Measure cost-effectiveness.**
- **Provide interpretive budget justification.**

What

Deciding what should be evaluated is the next component of determining the appropriate methods for conducting assessment. Numerous aspects of the interpretive services provided in your park can and should be evaluated, such as facilities, accessibility, interpreters, programs, information delivery services, publications, exhibits, etc. However, in keeping with the purpose of this handbook, we will exclusively focus on *personal* interpretive service evaluation.

The three categories of personal interpretive services relevant to evaluation are the interpreter, the audience, and the program. As discussed in Module 3-Communications each of these elements acts and reacts with the others to create the "whole" of the interpretive experience. Thorough evaluations must consider each of these elements to gain a true picture. For example, the interpreter could be a fabulous communicator and entertain the audience very well. However, without a meaningful message, the "whole" of the interpretive service would fall short. Conversely, the program could be technically "accurate" but be performed in a monotone by a lifeless interpreter. Once again, the "whole" of the interpretive service would be less than desired. As a result, when conducting evaluation, it is important to examine the elements from numerous perspectives. We will return to this discussion later in the module.

What should be evaluated?

Program

Relevant
Accurate
Provocative/enjoyable
Programmatically accessible
Organized
Retained
Thematic
Cost-effective

Interpreter

Body language
Appearance
Enthusiasm
Credibility
Subject matter
Voice quality
Eye contact
Confidence
Grammar

Audience

Learning
Attendance
Satisfaction
Enjoyment
Behavior change
Emotional impact
Attention
Memory

Interpreter

When referring to an interpreter in terms of evaluation, all who deliver interpretive services to the public are included. For example, volunteers, docents (highly trained volunteers), seasonal interpreters, etc., should all be evaluated. The DPR 461 series of forms are methods of assessing the interpreter. The interpreter controls program delivery and content and interacts directly with the audience. Examining the interpreter is a critical component of all evaluation of interpretive services.

Evaluations of the interpreter should include verbal and nonverbal communication techniques and skills, interpersonal interactions with the audience, and expertise in and ability to address RAPPORT elements (see Module 3-Communications). Interpreters can be evaluated using numerous methods (surveys, observations, etc.) implemented by a variety of sources (supervisors, audience, etc.). We will review specific methods later in this module.

Continuous assessment assures quality interpretation.
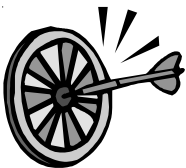
Audience

Another focus of evaluation is the audience. This might be a group of school children, kids at a Junior Ranger Program, a mixed-age audience at a campfire program, or general visitors encountered during roving. Anyone who receives interpretive services is the audience. Attendance, attention, satisfaction, and recall are some of the most common elements evaluated. CSP has standardized the collection of many of these elements in a statewide visitor satisfaction survey, the Visitor RAPPORT Survey (DPR 461A), and a School Group Program Evaluation form. Each form is designed to target selective audiences and collect specific types of information. We will discuss this in more detail later in the module.

Care should always be taken when collecting data to ensure the information will be helpful. For example, how many people attend the program is probably the most common form of data collected. It is easy to assess and gives some indication of the program's success, does it not? Not necessarily.

Although attendance level data is useful, popular programs do not necessarily equal effective programs. A critical question of evaluation should be, *"Is the program effective?"* Who better to answer that question than the audience? There are numerous ways that the audience can evaluate interpretive services. Since satisfaction, program objectives, and overall success of interpretation are judged based on the target audience, audience evaluations are a critical component of interpretive assessment. Just be sure you know why you are asking the question.

Program

The third target for evaluation is the program itself. Whether it is a formal school-based program or an informal roving contact, evaluations must include an examination of the program. There are many elements regarding a program that can be evaluated: visitor accessibility, meeting goals and objectives, relevance, organization, cost-effectiveness, etc.

As stated earlier, the elements of evaluation are interrelated. For example, the interpreter delivers the program and must be considered in the program evaluation through assessment of the delivery characteristics. The program should have specific, measurable objectives that reflect careful planning. These objectives should clearly answer the questions: "What is the goal of the program?" and "How should effectiveness be measured?" In other words, the objectives of the program determine the direction of the evaluation. The program is the focal point of the evaluation, but elements surrounding the effectiveness of the interpreter, the satisfaction of the audience, and meeting goals and objectives of the program should all be included in the assessment. The key is to thoughtfully consider how the audience, the interpreter, and the program evaluations impact each other.

When

"When should evaluation be conducted" is another question that must be answered in order to determine the best methods for assessment. There are three basic times to conduct evaluation: before, during, and after the program. Later in this module, we will discuss how to incorporate all three in an evaluation strategy.

Before (front-end)

Evaluation conducted before the program is called front-end, and it occurs during the planning stages. It involves research and careful planning to select the topic, theme, target audience, method of delivery, and goals/objectives of the program. Front-end evaluation is similar to a needs assessment where target audience needs, wants, and goals are examined and programs are matched accordingly (DeGraff, et al., 1999). This stage of the process is critical for evaluating the program in the later stages. It is in this front-end planning stage that the direction, goals, and objectives for the program are established. These goals and objectives, written in WAMS style (see Module 4-Planning), are the very indicators used to evaluate effectiveness in the later stages of evaluation. Module 4-Planning reviews in detail the elements of effective front-end evaluations.

During (formative)

   Formative evaluation occurs during the program and is often used by interpreters to ascertain if the program is going well. It occurs naturally when we watch body language, facial expressions, and other cues that indicate whether our communication is effective. Visitors' eye contact and questions are frequently used for formative assessment. Although these are not very scientific, they are classic examples of assessing interpretation during the program.

After (summative)

   Depending on your perspective, summative evaluation could occur at the end of an individual program or after an entire interpretive season. For example, a manager might consider all evaluations occurring during the season to be formative and only end-of-the-year, cumulative assessments to be summative. The DPR 918 is the semiannual interpretive summary report.

   Summative evaluation conducted at the end of an individual program is beneficial for discovering the impact of a program. Impact assessment (product evaluation) is considered one of the most important forms of assessment (Diamond, 1999; DeGraff, et al., 1999). Program effectiveness can be partially judged based on whether or not the outcomes matched the objectives of the program. If the answer is no, modifications to the objectives, the program, the target audience, the interpreter, or the program are in order. Thorough evaluation will make it clear which should be modified.

Closing the loop (modification)

   Modification to the original program, if warranted, is the next step of evaluation. This highlights one of the key evaluative elements, namely that it is a process and should always influence and inform subsequent steps. Reflecting back on Module 4-Planning you will recall that we discussed how monitoring and evaluations are an integral part of the creation process for interpretive programs. A report produced from the evaluation is not the final step; instead it is part of a cyclical process returning to its origin. Evaluations can suggest program modifications, give statistical summaries, and provide information for year-end reports regarding the products, successes, and satisfactions with interpretive efforts.

   Evaluations provide results that may not have been expected. For example, in *Making the Grade* (2002), an analysis of California State Parks' school group programs (1998-2002), one of the primary conclusions was that the evaluation forms for school groups should be modified. These results, although not the target information sought from the survey, serve to improve the evaluation process in the future and yield more programmatically useful data. Evaluations should be an integral part of programs, subsequent assessments, and the resulting modifications.

Putting it all together: an example

Front-end (before)
- Goals and objectives for a program are outlined in the planning process.
- A target audience for the program is determined based on research.
- Messages and themes are identified to meet audience needs and fulfill goals and objectives.

Formative (during)
- Interpreter communication characteristics during the program are assessed.
- Audience reactions to the program are examined.
- Messages and themes are analyzed to determine if appropriate for audience needs.

Summative (after)
- Message retention and recall by the audience are examined.
- Interpreter self-assessments are conducted.
- Behavioral impact of the program is assessed.

Closing the loop (modification)
- Modifications to program, delivery, and assessment methods are made.
- Program goals and objectives are reevaluated.

Who

The person or group providing the evaluation clearly affects how the process is conducted. The audience, supervisor, peer, expert, team, and the interpreter are all potential sources of evaluation. Each of these evaluators may use different methods and strategies of assessment. The following section provides an overview of all of the primary sources of evaluation to answer our question, *"Who is conducting the evaluation?"*

Audience

The audience members certainly provide feedback and evaluation of interpretation through their body language, questions, and written and verbal comments. The audience may be especially effective at telling us if program goals and objectives were met, if their needs were addressed, if the program itself was effective, and what their overall satisfaction level was. This feedback is usually initiated and facilitated by someone other than the audience, such as the interpreter, supervisor, expert evaluator, etc. Although anyone can initiate evaluations, it is your responsibility to ensure that it occurs. *Aiming for Excellence* offers an assessment tool (DPR 461A) that permits the audience to evaluate both the interpreter and the program. The quality of the feedback is only as good as the methods used to collect it. The DPR 461A, Visitor RAPPORT Survey, is a useful tool that may provide helpful data to improve your services. We will discuss other useful and practical methods later in this module.

Supervisor

A supervisor or Field Training Officer (FTO) assessment is very beneficial. This type of evaluation is often more objective than that provided by an audience. It can provide insights into the program's effectiveness, the interpreter's skills and abilities, audience reactions, and how well the program meets Department and program goals and objectives. The DPR 461 provides a framework for this type of evaluation by following the standard RAPPORT guidelines closely. Supervisory evaluation typically occurs at specific intervals throughout an interpretive season and is often tied to the systematic criteria outlined in *Aiming for Excellence*. However, supervisory evaluations are tied to components other than improving program effectiveness and can be intimidating. Nevertheless, it is in your best interest to seek out opportunities for your supervisor to evaluate you. Your supervisor will discuss the results with you, provide a written evaluation, and issue a followup report. *Aiming for Excellence* will prepare you for the specific RAPPORT criteria that your supervisor will be addressing. Supervisory evaluations are mainly geared toward coaching you through the developmental stages of your interpretive career. Use the feedback and coaching from supervisors to help improve your interpretive techniques, skills, and abilities.

Peer

Peer evaluation occurs when a colleague conducts the evaluation. Useful on many levels and highly recommended, these evaluations can increase overall effectiveness, build collegiality, enhance comprehensive knowledge of the resource, and provide nonthreatening feedback. It is important that the parameters of the evaluation are predetermined. Discuss ahead of time what you expect, establish a "game plan," and whether you would like oral and/or written comments. Peer evaluations can be accomplished using the DPR 461. Peers can also attend the program as members of the general audience and observe visitor reactions and comments throughout the program. This unsolicited audience feedback can provide authentic insight into program and interpreter effectiveness. But be sure to provide specific criteria for your peer to examine.

Expert

Professional evaluators, researchers, or statisticians perform expert and objective evaluations. These professionals have the advantage of being neutral and often are more able to recognize problems or issues that may be overlooked by peers or supervisors. There are many methods and strategies for expert evaluation, including in-depth interviews, focus groups, observations, and surveys. Expert evaluations are typically more costly and time consuming than other types of evaluations. However, the precision and quality of the information is often worth the cost. Following the recommendations of *Aiming for Excellence*, be sure you clearly identify what the goals and objectives are for the evaluation and how much time, money, and staffing are available. Research carefully to find the best expert for your needs, budget, and time frame. Often your District Interpretive Specialist can provide expert evaluations.

Team

Team evaluation is another form of assessment conducted in California State Parks. In fact, "It is recommended that each district form a District Interpretive Improvement Team (DIIT)…whose primary role is to facilitate the ongoing improvement of the district's entire interpretive program." DIIT evaluation is beneficial for its broad range of opinions and perspectives, the collegiality it builds, and the enhancement it encourages for conducting mission-driven interpretation. Evaluation teams "may choose an alternate name that is more specific to the programs of the district or that better serves to express the purpose of the team" (McDonald, 2000, p. 5).

> ***…every performance, every summer season, every service, and the entire program merits some kind of serious, systematic, open, fair analysis…***
>
> Douglas Knudson, Ted Cable, and Larry Beck

Interpreter

Interpreters themselves must also initiate evaluations. As the interpreter, you have the primary responsibility to evaluate your programs, yourself, and audience reactions to and satisfaction with the interpretation. Evaluations conducted by interpreters of themselves are called self-evaluation. Knudson, Cable, and Beck (1995) said this is the "most efficient and least threatening method of evaluation" (p. 448). *Aiming for Excellence* provides a self-evaluation form (DPR 461D) using the RAPPORT criteria.

The remaining sections of this module will review and discuss practical and effective methods that you can implement in order to evaluate the effectiveness of personal interpretive services. There are disadvantages and advantages to evaluating your own programs. We will review and discuss these as we examine the various methods for conducting evaluations.

Each district should have a District Interpretive Improvement Team (DIIT).

How

Traditional and scientific

There are numerous methods for how evaluations should be conducted, ranging from rigorous scientific surveys to casual observations of visitors' behavior. In fact, most authors in literature make two classic distinctions of assessment methods: traditional and scientific. There have been many debates regarding the appropriateness of rigorous scientific assessment versus the more traditional, gut-level impressions about programs. This is similar to the discussion of whether or not interpretation is an art or a science. Many interpreters agree with William Lewis that "after having observed, practiced, and coached interpretation in parks for more than thirty years, that the gut-level type of evaluation is satisfactory for more purposes" (Lewis, 1986, p. 110). Seeing the sparkle in the eyes of visitors and sensing the enthusiasm in the audience are both "traditional" methods of evaluation. Interpreters who have worked for any length of time can tell you when their programs went well. These "traditional" proponents view scientific evaluation as minimizing the true impact and effect of interpretation on the heart and mind of the individual. After all, we are exposing the "soul of a thing" (Tilden, 1977) to the heart of a person. How can this be measured, and does not the mere attempt to assess it degrade the spirituality of it? How can you measure and quantify art?

Interpreters that align more closely with the scientific forms of evaluation agree that, "current emphasis on accountability means interpretation must justify the time, money, equipment, and personnel it requires" (Sharpe, G., and Sharpe, W., 1986, p. 97). This position asserts that interpretation may be spiritual in nature, but it must also be accountable to the people it serves, the agency it represents, and the resources it interprets. Just as ministers or priests may be "called" to the church, they still try to accomplish specific goals and are periodically evaluated on their effectiveness in meeting those goals. It does not lessen their spirituality, but it is an indicator of their commitment and dedication to it.

Scientific evaluation provides defensible results with rigorous, tangible outcomes as opposed to the more traditional methods of assessment. For example, your gut-level feeling about your program is not easily demonstrated to a supervisor nor is it an adequate justification for an increase in budget. Scientific assessment is a more accurate and representative measure of effectiveness. It is the position of California State Parks that evaluation be conducted in a systematic, consistent manner. "Interpretive services in California State Parks play a key role in the organization's mission - inspiring and educating the people of California and creating opportunities for high-quality outdoor recreation. Measurement of the success of that mission is critical to maintaining support for interpretive services . Therefore, it is imperative that evaluations take place in a consistent and systematic way" (Aiming for Excellence, p. vii).

> *The more we live by our intellect, the less we understand the meaning of life.*
>
> Leo Tolstoy

Scientific evaluation

**Questionnaires, formal/informal interviews, focus groups, expert observations**

Pros	*Cons*
• Generates definitive evidence of effectiveness	• Requires money, time, and expertise
• Respected by researchers and administrators	• Difficult to collect, tabulate, and analyze
• Results in quantifiable data	• Unable to quantify intangibles
• Allows for precise measurement of interpretive services	• Field application of results is often difficult

Since an objective of any park administration is to improve the quality of park use, the effectiveness of our interpretive program is a major concern of all administrators.

George Hartzog
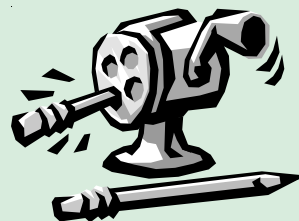
Traditional evaluation

Visitor body language, eye contact, attention, attendance, and questions

Pros	*Cons*
• Provides immediate feedback	• Often inaccurate and unreliable
• Easy to implement	• Biased
• Requires little training	• Based on emotion and feeling
• Based on intuition and gut feeling of interpreter	• Provides no distinction between entertainment and interpretation

There are clearly merits to each evaluative approach and a time and place for each method. Both methods should be used to gain the most accurate "total" picture of the effectiveness of interpretation in your park. As we have just discussed, the most appropriate methods for conducting an evaluation depend on the answers to many questions such as: Who is conducting the evaluation? What is being evaluated? When will the evaluation take place? They also depend on many other issues, such as money, time, and the expertise available. Let us proceed with the practical realization that you, as a park interpreter, have no money, know very little about evaluation, and are already feeling overwhelmed with the amount of work you have to do. So how can you possibly conduct quality evaluation?

Objective-based assessment

Before beginning a discussion of specific evaluative methods, we need to establish a framework through which to review them. For example, in discussing surveys, we could focus on who is conducting the survey (audience, supervisor, peer, etc), when it would be conducted (front-end, formative, etc.), and/or what it will examine (interpreter, program effectiveness, meeting goals and objectives, etc.). Since each of the guiding questions introduced at the start of this module impacts and affects the others, let us return to one of the guiding issues underlying all interpretation: What is the objective of the program? As discussed in Module 4-Planning, all programs should have clearly stated objectives. Methods of evaluation, regardless of when, by whom, or where they are conducted, should be based upon assessments of what you were trying to accomplish with programs (Knudson, et al., 1995; Diamond, 1999). This idea promotes a return to that original question posed in Module 2-Purpose and Values-Why do we conduct interpretation?

The purpose behind the program should always be clear. In addition, good objectives should demonstrate the connections among front-end, formative, and summative assessment. We have seen in Module 4-Planning there are three basic types of objectives for interpretive programs: cognitive, behavioral, and affective. Let us use these objectives to guide our discussion of the associated methods of assessment for each type of objective.

Cognitive objectives

Cognitive objectives involve the understanding, information transfer, and/or learning that occurs based on the program. An example of this would be: *During a quiz at the end of the program, 80 percent of randomly selected visitors will correctly identify two reasons to protect tide pools*. One of the classic methods of measuring cognitive objectives is through pre/post assessment. Knowledge is measured before and after exposure to a program to determine if a difference exists. If the evaluations are conducted appropriately, the increase in knowledge after exposure to the program can be attributed to the program. This information can be used to document meeting (or not meeting) the objectives. This is one of the most common forms of assessment, primarily due to the ease of data collection and the tangible results produced. The DPR 461 (Standard RAPPORT Evaluation), DPR 461A (Visitor RAPPORT Survey), and School Group Program Evaluation each provide some questions that assess cognitive aspects of the program. The box on the following page presents some additional, easy techniques that you can implement to demonstrate program effectiveness in meeting cognitive objectives.
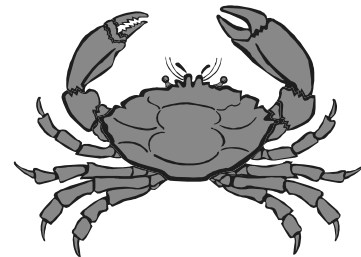
> *...we need to spend more time developing program objectives. Objectives are an essential precursor to evaluation. Without carefully written objectives, it is next to impossible to interpret the results of evaluation.*
>
> Maureen McDonough

Cognitive objective assessment techniques

- **Oral quizzes–**Learning can be easily assessed before, during, or after a program by having a quiz. Quizzes provide enjoyment, hands-on participation and social interaction, and allow for assessment. Divide visitors into teams, or have them participate as individuals. Postcards, pencils, or other appropriate items can be used as "prizes." Quizzes may involve answers to questions or identification of objects, smells, etc. Use your imagination to make the evaluation part of the program fun.

- **Pre/post written responses–**This is a very useful method to assess cognitive objectives. Have two colors (for our example, yellow and green) of 3 x 5 index cards, and hand out one card of each color and a pencil to each individual at the start of the program. After your introduction, ask visitors to put the last three digits of their social security number in the top left corner of both cards. This is so you can match the two cards later. Have them write the answer to a question(s) you ask at the start of the program on the green card and collect them. Conduct the program as planned, and repeat the question(s) at the end of the program. Have them record their responses on the other card. After the program, the cards can be matched, using the numbers, to assess any increase in knowledge. Note: If there are children in the group too young to meaningfully participate, have them draw a picture of something from the program before and after. In addition, letting visitors keep the pencil is a nice touch and helps build public support and agency recognition.

- **Review questions–**Similar to the above methods is the use of review questions during and after the program. Ask questions to assess visitors' knowledge levels of information presented during the program. These can be conducted formally or informally as the program proceeds. For example, on the return after a guided walk, ask visitors questions about information covered during the walk.

- **Response or comment cards–**Response or comment cards (see box on the next page) are not only beneficial for measuring cognitive objectives but effectiveness as well. *Aiming for Excellence* has several good examples. Response cards are designed to measure how well the program meets performance objectives. They are easy, convenient, and effective measures that first-time interpreters can implement.

Prairie Creek Redwoods State Park

Please help us improve the quality of our interpretation.

1) What was the main point or message of the program?

2) How did you find out about the program?

3) What did you enjoy most about the program?

4) What did you enjoy least?

5) What is the Save-the-Redwoods League?

Thank you for your time!

Cognitive objective assessment tips

- **Randomly distribute response cards/ask survey questions–**For example, every fifth person, every other person, etc., produces results more reflective of the entire population of users. Do not ask too many questions, no more than five in any data collection method or from any visitor. Too much data will be difficult to tabulate and analyze. In addition, visitors are on vacation and should not be needlessly disrupted from their recreational experience.
- **Multiple methods of return–**Provide multiple drop-off points for any written data. Make it easy and anonymous for visitors to return the information.
- **Assess throughout the season–**Conduct evaluations throughout the interpretive season. This information can be used very effectively to show patterns and changes.
- **If written, keep small–**Keep the actual written surface small and unimposing. This also forces you to carefully select the questions.
- **If oral, have a method of recording–**Have a coworker or friend available to discreetly record answers to any oral methods of assessment. If that is not possible, "placeholders" can be used to serve as hatch marks of answers. (For example, move coins from one pocket to another in between contacts to signify the number of correct responses.)
- **Pre/post assessment (change)–**Incorporate before and after assessment methods to determine the cognitive impact of programs.
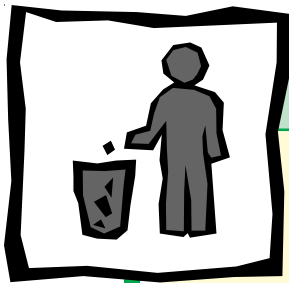
Behavioral objectives

Behavioral objectives focus on the physical behaviors of visitors. For example, "50 percent of visitors will take a brochure on tide pool protection," is a behavioral objective. One of the classic methods of measuring this type of objective is by observation. The key to implementing and measuring behavioral objectives is to ensure that observation is easy, practical, and discreet, and can reveal the targeted behavior. For example, if the desired behavior is to "protect" the resource, then what behaviors will you tangibly measure? Objectives must be stated in terms of observable phenomena. One of the most common problems with using behavioral objectives to measure program success is that objectives often attempt to assess behaviors that are not easily observed or connected with program exposure. The box on the next page presents some simple tips for conducting evaluation based on behavioral objectives.

Behavioral objective: 50% will take a brochure.

Behavioral objective assessments: an example

If the behavioral objective in your program is to change visitors' actions regarding littering the park, then observing visitors' responses to a "planted" piece of garbage may be indicative of program success. For example, before the program, observe random visitors in the parking lot and their response to a "planted" piece of trash. How many visitors walk by the trash before someone picks it up? If 40 visitors walk by before someone picks it up, this might be a general average for this behavior–one out of 40 visitors will pick up trash. Repeat this several times to get a true average for the behavior. After the program, count how many visitors walk by before someone picks up the same type of "planted" trash. If only 10 program attendees walk by the garbage before someone picks it up, then the program may have reduced the average to one out of 10 visitors. This produces a tangible behavioral result from the program.

Keep in mind that this is not statistically defensible data. However, it does yield practical results that if found repeatedly or averaged over time can produce meaningful results of our interpretive efforts.

Behavioral objective assessment tips

- **Conduct random observations of visitors–**Randomly choose times, places, and visitors, such as every fifth person, every other day, every other parking lot, etc. This produces results more reflective of the general population of park users.
- **Implement pre/post observations–**Observations of random visitors in the park can be compared to random observations of visitors exposed to the program to determine a "difference" due to exposure.
- **Target only specific observations–**Keep observations very specific and manageable. Do not conduct observations of "resource protection." These observations are often too general to be very useful. Instead, determine what specific behaviors reflect resource protection.
- **Assess behavior during the program–**Conduct observations during or immediately after the program. Assessments should be linked as closely as possible to the time of exposure to the program.
- **Observe evidence of behavior–**Sometimes it is easier and more discreet to observe the *evidence* of the behavior. For example, instead of watching visitors pick up trash, count the number of pieces of planted litter along a trail before and after the program.
- **Be sure the behavior is observable–**If the desired behavior is to "provoke" visitors to want to learn more, then you must be able to observe and assess that behavior. How do you think this could be done?

Affective objectives

Affective objectives focus on the emotional impact of programs. The goal is for visitors to care, to support, to feel, and to be changed by the emotional impact of the resource and our programs. This is one of the most powerful types of program objectives, but also one of the most difficult to assess. "Visitors randomly surveyed before the program compared to those surveyed after will demonstrate a 20 percent increase in positive feelings that tide pools should be protected," is an example of an affective objective. One of the classic methods of measuring this type of an objective is a self-report of emotion. The Visitor RAPPORT Survey (DPR 461A) is an example of allowing visitors to report on how they feel about programs. Observations can also be used to measure targeted affective objectives. For example, the number of visitors purchasing books or taking brochures about the program topic after exposure to it could assess whether visitors care about the topic. Keep it simple and be sure that observations or self-report measures reflect the targeted emotions. Tips and techniques for conducting this type of evaluation are similar to the strategies we reviewed for cognitive and behavioral objectives. For example, if using self-report measures of emotions or feelings, follow the tips presented on page 12-14. If using behavior to measure effectiveness, the tips above will be helpful.
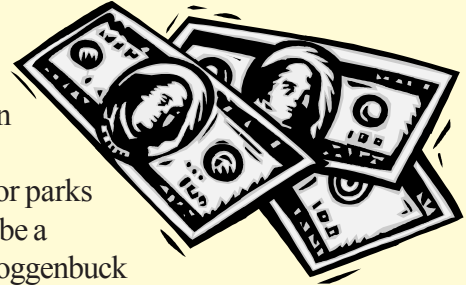
Cost effectiveness

Another useful element to include is a measure of cost effectiveness. Although it is not the primary focus for determining the program's effectiveness, it is useful information to include in reports targeted at managers.

Like it or not, budgets drive much of the practical ability for parks to conduct interpretation. Determining cost effectiveness can be a significant contributor when providing budget justifications (Roggenbuck and Propst, 1981; Machlis, 1986). However, assigning a dollar figure to the worth of interpretation is a double-edged sword. Although it often provides a tangible measure, many would argue that you cannot place a dollar value on intangibles, such as inspiration or provocation. Although we agree, you can place a value on some of the outputs. For example, cost effectiveness can be measured in the reduction of litter after a program and the budgetary savings of reduced cleanup hours spent by personnel. In addition, an increase in book sales after a program on the same topic can provide a good measure of cost benefit.

In addition to measuring output, the cost effectiveness of a program can also be evaluated by assessing input. For example, the hours spent preparing and delivering a program can be converted into the equivalent amount of salary (P) and then divided by the total number of people contacted in the program (N) to ascertain the cost per individual visitor (P/N) of the program. This measure, coupled with any discernible outputs in dollar figures, can provide a compelling picture of the cost effectiveness of a program.

Putting it all together

The best solution for increasing the overall effectiveness of evaluation is to use several methods. Because each method has strengths and weaknesses associated with it, evaluation through multiple assessment strategies will produce a more accurate assessment of program effectiveness. This is called "triangulation." Triangulation is used to increase overall accuracy. For example, if three separate methods are implemented to assess meeting program objectives and all three produce similar results, this is a clearer result than if only one method was used. Methods discussed in this module are often not as rigorous as the methods used by researchers or expert evaluators. However, triangulation increases the overall accuracy of the evaluation. Different methods can be used within one program or across the season of the program to produce a summative evaluation. Regardless of the approach, increased accuracy will be achieved if multiple methods are used (see box on next page).

Evaluation strategy - an example

Program

Afternoon tide pool walk.

Th eme

Tide pool life strives to survive with each changing tide.

Goals

1. Increase visitor satisfaction with the experience of exploring tide pools.
2. Promote tide pool protection.
3. Increase visitor knowledge about the fragile nature of tide pools.

Objectives

1. Visitors surveyed randomly before and after the program will demonstrate a 20 percent increase in positive feelings that tide pools should be protected.
2. 70 percent of randomly surveyed visitors responding to the general park satisfaction survey will indicate positive satisfaction levels with the program.
3. 80 percent of visitors asked will correctly identify two reasons to protect tide pools.

The following methods would be implemented throughout the entire season.

Audience

1. Have randomly selected visitors complete comment cards regarding tide pool objectives.
2. Implement the Visitor RAPPORT Survey (DPR 461A).
3. Conduct visitor satisfaction surveys.

Peer

1. Observe visitor behavior and comments during and after the program, specifically examining protection behaviors and overall satisfaction comments.
2. Use/complete Standard RAPPORT Evaluation (DPR 461).

Supervisor

1. Use/complete Standard RAPPORT Evaluation (DPR 461).
2. Invite supervisors to your program early in the season, the middle, and near the end so they can see the improvements made based on their feedback.

Interpreter

1. Conduct self-evaluations throughout the season using Self-Evaluation of Interpretive Programs (DPR 461D).
2. Observe visitor behavior at the tide pools before and after your programs.
3. Collect attendance levels at all programs.

This outline provides a strategy for evaluating the program that reflects many different points of view, times throughout the season, and purposes for the evaluation. By conducting an evaluation that relies on more than one method, you gain greater insight into the true success of your program. No one method is perfect, but together several different ones can provide a better picture of the interpretive services you provide.

As you can see in the box on the previous page, RAPPORT evaluation forms provide a method of "pulling it all together." Because they are based on the primary criteria of effective interpretation (Module 3-Communications), the DPR 461 series helps to measure a program's success on many levels. In addition, the series can be used by interpreters, supervisors, peers, and the audience. They expose overall visitor satisfaction, meet visitor needs and objectives, and assess program and interpreter effectiveness.

Reporting

Another aspect of conducting evaluation is documenting the results for supervisors and administrators. Reporting should be simple, brief, and directly related to the documented information. Remember, this process is not meant to replace the existing reporting procedures outlined in *Aiming for Excellence*, but only to serve as a supplement. The primary use of this information is to help you, the interpreter, increase effectiveness. However, since you are collecting the information, it may as well demonstrate your effectiveness to others. Reporting should include program summaries, outlines, objectives, attendance levels, and summaries of assessment measures used. For example, you could simply include an outline of the box on page 12-18 and the results from each of the methods used throughout the season.

Be sure to keep a copy of all your reports and reporting procedures for your own records. Do not forget to place a copy of the reports in the individual program folders created in the planning stage (see Module 3-Communications). This information will be very useful to future interpreters.

Evaluation is an extremely useful tool to indicate program effectiveness, encourage interpreter development and growth, and further the mission of the Department. When asked why you are conducting an interpretive program, you should now be able to answer and clearly demonstrate your effectiveness in meeting those goals.

What's ahead

We have completed our journey in understanding what interpretation is, how to conduct various program types, and finally how to evaluate the success of our efforts. Now we turn to the profession and its future. What does it mean to be an interpreter? What are the responsibilities of the profession, and what role do interpreters play to ensure further development of the discipline? In Module 13-Professionalism we will discuss these and other issues facing the profession of interpretation.

Brochu, Lisa, and Tim Merriman. *Interpretive Guide Training Workbook.* Fort Collins, CO: National Association for Interpretation, 2000.

Brochu, Lisa, and Tim Merriman. *Personal Interpretation.* Fort Collins, CO: National Association for Interpretation, 2002.

Saunders, Jenan. *Making the Grade: An Analysis of California State Parks' School Group Program Evaluation 1998-2002.* Sacramento, CA: California Department of Parks and Recreation, 2002.

DeGraaf, Donald, Debra Jordan, and Kathy DeGraff. *Programming for Parks, Recreation, and Leisure Services: A Servant Leadership Approach.* State College, PA: Venture Publishing, Inc., 1999.

Diamond, Judy. *Practical Evaluation Guide: Tools for Museums & Other Informal Educational Settings.* Walnut Creek, CA: Alta Mira Press, 1999.

Huggins, R. "Going for the Gold". *Interpretive Views,* edited by G. Machlis. Washington, DC: National Parks and Conservation Association, 1986: 65-70.

Knudson, Douglas, Ted Cable, and Larry Beck. *Interpretation of Cultural and Natural Resources.* State College, PA: Venture Publishing, Inc., 1995.

Lewis, L. "Evaluation of Interpretation". *Interpretive Views,* edited by G. Machlis. Washington, DC: National Parks and Conservation Association, 1986: 107-112.

Machlis, Gary. *Interpretive Views.* Washington, DC: National Parks and Conservation Association, 1986: 95-97.

McDonald, Linda. *Aiming for Excellence: An Evaluation Handbook for Interpretive Services in California State Parks.* Sacramento, CA: California Department of Parks and Recreation, 2000.

Roggenbuck, Joseph, and D. Propst. "Evaluation of Interpretation." *Journal of Interpretation* 6, no.1 (1981): 13-23.

Sharpe, Grant, and Wenonah Sharpe. "Evaluating the Role of Interpretation: *Interpretive Views,* edited by G. Machlis. Washington, DC: National Parks and Conservation Association, 1986: 95-97.

Tilden, Freeman. *Interpreting Our Heritage.* 3rd ed. Chapel Hill, NC: University of North Carolina Press, 1977.

Babbie, Earl. *The Practice of Social Science Research.* 6th ed. Belmont, CA: Wadsworth Publishing Company, 1992.
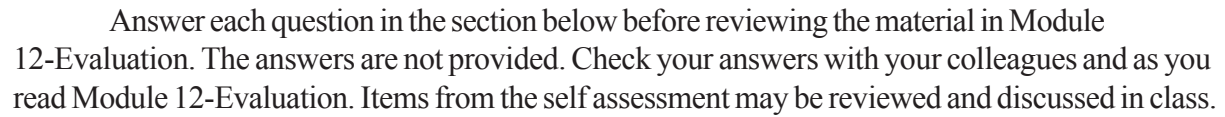
Henderson, Karla. *Dimensions of Choice: A Qualitative Approach to Recreation, Parks, and Leisure Research.* State College, PA: Venture Publishing, Inc., 1991.

Mitra, Amanda, and Sam Lankford. *Research Methods in Park, Recreation, and Leisure Services.* Champaign, IL: Sagamore Publishing, 1991.

"Interpretation as Communication." *Trends* 34, No.4 (1997). National Park Service.

Evaluation

Self assessment

Answer each question in the section below before reviewing the material in Module 12-Evaluation. The answers are not provided. Check your answers with your colleagues and as you read Module 12-Evaluation. Items from the self assessment may be reviewed and discussed in class.

1) What is *Aiming for Excellence*?

2) List four reasons why we should conduct evaluation of interpretive programs.

3) When conducting evaluation of personal interpretive services, what specific types of things should be evaluated?

4) What is the DPR 461?

5) When should evaluation be conducted?

6) RAPPORT criteria form the basis of much of the evaluation in California State Parks. (Explain your answer.)
   a) True
   b) False

7) Name two benefits of having an expert provide the evaluation.

8) Evaluation of interpretive performances should be conducted by a supervisor. (Explain your answer.)
   a) True
   b) False

9) What is self evaluation?

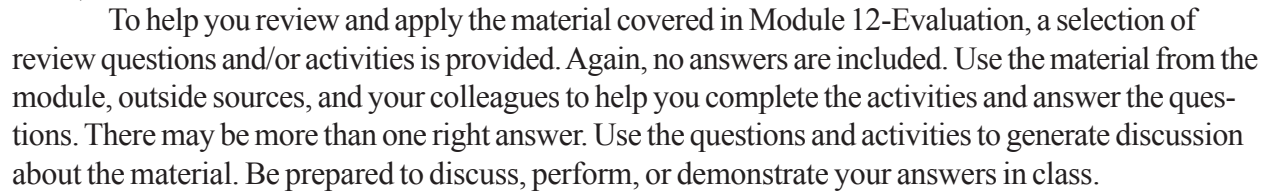10) Describe one method to assess learning based on program objectives.

11) To save time, you should only distribute surveys to visitors who you think will fill them out. (Explain your answer.)
   a) True
   b) False

12) You should have only one drop-off point for surveys. (Explain your answer.)
   a) True
   b) False

13) When measuring the behavioral impact of a program, which of the following are recommended practices? (Check all that apply.)
   a) Assess general, not specific, behavior
   b) Conduct random observations of visitors
   c) Observe evidence of behaviors
   d) Practice discreet observations

14) It is impossible to assess the emotional impact of a program. (Explain your answer.)
   a) True
   b) False

15) What is triangulation?

   Now that you have completed the self assessment questions, review the material in Module 12-Evaluation to confirm your answers. After reading the module, move on to the workbook learning activities, which will assist you in developing your skills.

To help you review and apply the material covered in Module 12-Evaluation, a selection of review questions and/or activities is provided. Again, no answers are included. Use the material from the module, outside sources, and your colleagues to help you complete the activities and answer the questions. There may be more than one right answer. Use the questions and activities to generate discussion about the material. Be prepared to discuss, perform, or demonstrate your answers in class.

1) You designed an interpretive program to educate visitors about two primary reasons to protect the resources in the park. Describe two methods you can use to demonstrate your effectiveness in meeting your program goals. Describe an additional method you could use to assess your effectiveness as an interpreter for this program.

2) A common goal for interpretation is provocation. Describe one way you can observe and measure if your program was provocative.

3) What do you think is the most important reason to conduct evaluation in your park?

4) You are busy and have many duties to perform in the course of a day. What steps can you take to ensure evaluation will not fall by the wayside?

Answer each question with the information specific to your park. You will have to conduct some research in order to answer each question. Use the answers as a guide for beginning your career in California State Parks.

Evaluation

Park name_____

1) Describe how you can use triangulation to evaluate a program you designed in one of the previous modules.

2) Look at the box on page 12-18 of the *Handbook*. Create an evaluation strategy for a program you are designing in your park. Be sure that it contains front-end, formative, and summative elements.